

misspellings, e.g., “f” may be substituted for “ph”, as in “fantom/phantom”, “Johnston” may be misspelled as “Johnson”, etc.

[0023] Embodiments of the present invention build on techniques, systems and methods disclosed in earlier filed applications, including but not limited to U.S. patent application Ser. No. 11/235,928, filed Sep. 27, 2005, entitled “Method and System for Processing Ambiguous, Multi-term Search Queries” and U.S. patent application Ser. No. 11/312,908, filed Dec. 20, 2005, entitled “Method and System for Dynamically Processing Ambiguous, Reduced Text Search Queries and Highlighting Results Thereof”; the contents of which are hereby incorporated by reference in their entirety. Those applications taught specific ways to incrementally process search queries containing search term prefixes formed from ambiguous text input. Similarly, embodiments presented herein may be used with the techniques, systems, and methods disclosed in earlier filed applications, including but not limited to U.S. patent application Ser. No. 11/356,788, filed Feb. 17, 2006, entitled “Method and System For Offsetting Network Latencies During Incremental Searching Using Local Caching and Predictive Fetching of Results From A Remote Server”; the contents of which are hereby incorporated by reference in their entirety. That application taught specific ways to retrieve and order content items for presentation to the user. The present techniques, however, are not limited to systems and methods disclosed in the incorporated patent applications. Thus, while reference to such systems and applications may be helpful, it is not believed necessary to understand the present embodiments or inventions.

[0024] FIG. 1 schematically illustrates an overall system for performing searches with reduced text entry using a wide range of devices in accordance with one or more embodiments of the invention. A server farm **101** can serve as the source of search data and relevance updates with a network **102** functioning as the distribution framework. The distribution framework could be a combination of wired and wireless connections. Examples of possible networks include cable television networks, satellite television networks, IP-based television networks, wireless and wired phone networks. The search devices could have a wide range of interface capabilities such as a hand-held device **103** (e.g., a telephone or PDA) with limited display size and a limited keypad, e.g., a keypad with overloaded keys, a small QWERTY keypad, or other keypad. The search devices could also include a television system **104a** coupled with a remote control device **104b** having a keypad with overloaded keys, a small QWERTY keypad, or other keypad. Finally, the search device could include a Personal Computer (PC) **105** with a full QWERTY or other keyboard and a computer display. FIG. A provides one example of a keypad having overloaded keys. In that example, keys corresponding to the numerals 2-9 have sets of alphanumeric characters associated with them.

[0025] FIG. 2 illustrates exemplary device configurations for performing searching in accordance with one or more embodiments of the invention. In one configuration, the television system **104a** has a display **201**, a processor **202**, volatile memory **203**, text input interface **204**, remote connectivity **205** to the server farm **101** through the network

102, and a persistent storage **206**. These elements may be contained in one device, or be connected through wired or wireless techniques **207**.

[0026] In another possible device configuration, the hand-held device **103** and television system **104a** might not have local persistent storage **206**. In such a configuration, the device can use remote connectivity **205** to submit the query to a server farm **101** and retrieve results from it.

[0027] In another exemplary configuration, the television system **104a** may not have remote connectivity **205**. In this configuration, the search database may be locally resident on a local persistent storage **206**. The persistent storage **206** may be a removable storage element such as SD, SmartMedia, CompactFlash card etc.

[0028] In a configuration of the television system with remote connectivity **205** and persistent storage **206** for searching, the device may use the remote connectivity for search relevance data update or for the case where the search database is distributed on the local storage **206** and on the server **101**.

[0029] In one or more embodiments of the invention, a television system **104a** may have a set-top box with a one-way link to a satellite. In this configuration, all search data including relevance updates may be downloaded to the device through the satellite link to perform local search. Relevance updates could be periodically done through this link.

[0030] FIG. 3 illustrates the pre-computation operations performed for orthographic and typographic error resilience, in one or more embodiments of the invention. This operation is performed on single and multiple descriptive terms constituting a phrase (e.g., “Armageddon”, “The guns of Navarone”, “Tom Cruise” and “Cruise Tom”) that can be used for item discovery in the search space. First, the method identifies ordered and unordered phrase and/or term combinations that describe the items of interest to the user (step **300**). Steps **301** and **302** apply orthographic and typographic rules on the terms constituting phrases to generate orthographic and typographic variants of the original terms. In an embodiment of the invention, the orthographic rules for generating variants of terms include (1) dropping of all vowels (e.g., “shylck pen” for “shylock pacino”), (2) using a SOUNDEX-like equivalence class for creating phonetically similar term variants (“fone” and “phone”), and (3) applying rules that further capture phonetic equivalence at specific character locations in a word. The orthographic rules can include those described in Patent Application No. 60/739,893, incorporated above, but embodiments of the invention are not limited to these rules.

[0031] Typographic variant generation rules include generating N-gram word fragments for each descriptive term or phrase. N-grams are incomplete portions of descriptive terms or phrases containing N number of characters, e.g., trigrams are a species of N-grams having three characters. The characters in the N-grams can be ordered according to the characters’ ordering in the term or phrase, but the characters can be non-sequential. In the illustrative embodiment of the invention, multiple trigrams are generated for each term by dropping characters from the term, e.g., “pci”, “pac”, “aci”, “acn”, “cin”, “cio”, etc. for “pacino”. N-grams may also be formed for a descriptive phrase. For example,